AD-A067 644 TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS F/G 12/1
MULTIPLE TIME SERIES MODELING II.(U)
FEB 79 E PARZEN, H J NEWTON N00014-78-C-0599
UNCLASSIFIED TR-N-1 NL

1 OF 1
AD
A067644

END
DATE
FILMED
6-79
DDC

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843

INSTITUTE OF STATISTICS
Phone 713 - 845-3141

MULTIPLE TIME SERIES MODELING II

By Emanuel Parzen and H. J. Newton
Institute of Statistics, Texas A&M University

Technical Report No. N-1

February 1979

Texas A & M Research Foundation
Project No. 3838

"Multiple Time Series Modeling and Time Series Theoretic Statistical Methods"

Sponsored by the Office of Naval Research

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. N-1
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) Multiple Time Series Modeling II
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Emanuel Parzen ~~and~~ H. J. Newton
8. CONTRACT OR GRANT NUMBER(s): (15) N00014-78-C-0599
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Texas A&M University, Institute of Statistics, College Station, TX 77843
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Code 436, Arlington, VA 22217
12. REPORT DATE: (11) Feb~~ruary~~ 1979
13. NUMBER OF PAGES: 61
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 64 p.
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report): NA (9) Technical rept.,
18. SUPPLEMENTARY NOTES: NA (14) TR-N-1
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Multiple time series, multiple spectral analysis, multiple autoregressive schemes, index time series, principal components of time series, canonical correlations of time series
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

This paper defines the problem of time series modeling as model identification (determining the predictor variables) and parameter identification (estimating the prediction filter and the prediction error covariance matrix). Various auto-regression and cross-regression representations are defined for a stationary multiple time series. The role of basic regression and latent value algorithms is discussed. It is suggested that principal component

next page

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

analysis of spectral density matrices may not be useful in practice, whereas autoregressive methods are. The problem of defining an index time series is discussed; an approach is described in terms of the notion of predictable components.

MULTIPLE TIME SERIES MODELING II

by

Emanuel Parzen and H. J. Newton*

Institute of Statistics
Texas A & M University

Abstract

This paper defines the problem of time series modeling as model identification (determining the predictor variables) and parameter identification (estimating the prediction filter and the prediction error covariance matrix). Various auto-regression and cross-regression representations are defined for a stationary multiple time series. The role of basic regression and latent value algorithms is discussed. It is suggested that principal component analysis of spectral density matrices may not be useful in practice, whereas autoregressive methods are. The problem of defining an index time series is discussed; an approach is described in terms of the notion of predictable components.

1. Introduction

The problem of multiple time series modeling (or equivalently, the problem of system identification) has an extensive theory; see, for example, Akaike (1976), Brillinger (1975), Mehra and Lainiotis (1976),

*Research supported by the Office of Naval Research.
(Contract N00014-78-C-0599)

Priestley (1976), (1978), Priestley, Subba Rao, and Tong (1973), and Subba Rao (1975). However, its practical implementation still seems to be at its beginning stages. This paper, a sequel to Parzen (1967), (1969), (1977), aims to describe some practical approaches to empirical modeling of multiple stationary time series which have been implemented in our extensive library of computer programs called TIMESBOARD.

The theory and practice of statistical data analysis involves answers to two kinds of questions:

(1) Parameter identification: what are the most likely parameter values (assuming a probability model for the data which is specified up to a finite number of parameters), and

(2) Model identification: what are models that adequately fit the observed sample probabilities, moments, order statistics (quantile functions), and correlations.

Empirical multiple time series analysis is concerned with finding relations among (normal) time series $Y_1(t), \ldots, Y_d(t)$, $X_1(t), \ldots, X_r(t)$, where Y_j is used to denote a "dependent" or "output" variable and X_k is used to denote an "independent" or "input" variable. In classical multivariate analysis observations at different times t are assumed independent and one can confine the search for relationship to relationship between contemporaneous variables. In time series analysis observations at different times t are correlated and a major goal is to identify the time lags at which significant relationships exist.

When a multiple time series $\{Y(t),\ t = 0, \pm1, \pm2, \ldots\}$ is assumed to be normal, stationary, and have zero means, its probability law is specified by its covariance matrix function

$$R(v) = E[Y(t)\, Y^*(t+v)], \quad v = 0, \pm1, \ldots \quad ;$$

we use the following notation on matrices: $*$ to denote complex conjugate transpose, $'$ transpose, and $\bar{\ }$ complex conjugate.

The assumption of zero means is to be interpreted that $Y(\cdot)$ has been pre-processed (by a detrending procedure, say) to approximately obey this assumption. One then says that $Y(\cdot)$ has been detrended.

2. Basic Regression and Latent Value Algorithms

Our approach to multiple time series modeling is based on using in suitable ways certain algorithms of conventional regression and multivariate analysis. Let X and Y be jointly normal complex-valued random vectors with zero means and partitioned covariance matrix

$$\Gamma = \begin{bmatrix} \Gamma_{XX} & \Gamma_{XY} \\ \Gamma_{YX} & \Gamma_{YY} \end{bmatrix} = E \begin{bmatrix} X \\ Y \end{bmatrix} \begin{bmatrix} X^* & Y^* \end{bmatrix}$$

It should be noted that in time series analysis Y could denote the values of a time series at the present time t , and X could denote the values of the time series at past times.

I. Regression of Y on X . One can express Y as a part Y^μ (linearly) explained by X and a residual $Y^\nu = Y - Y^\mu$. One seeks the regression coefficient matrix B such that

$$Y^\mu = E[Y|X] = BX$$

and one seeks the residual covariance matrix Σ such that

$$\Sigma = E[Y^\nu (Y^\nu)^*] \ .$$

Since $E[Y^\mu X^*] = E[Y X^*]$, by definition of conditional expectation, it follows that

$$B = \Gamma_{YX}\Gamma_{XX}^{-1}$$

$$\Sigma = \Gamma_{YY} - \Gamma_{YX}\Gamma_{XX}^{-1}\Gamma_{XY}$$

Using matrix pivoting procedures one can transform Γ to a partitioned matrix

$$\begin{bmatrix} \Gamma_{XX}^{-1} & B' \\ -B & \Sigma \end{bmatrix}$$

which may be the computationally most convenient way to compute B and If Γ is not known theoretically but is estimated from data one estimates B and Σ using the foregoing formulas with an estimator of Γ .

II. Subset Regression. A basic technique of model identification is reduction of the dimension of the input vector X , in order to find a "parsimonious" set of explanatory variables represented by the components of the vector X of as low dimension as possible. The techniques for doing this (called Subset Regression or regression model identification procedures) choose a "parsimonious" set of variables from among all possible subsets of components of X by computing, and comparing, the

Σ matrix for each possible definition of X. These procedures are easiest to interpret when Y is scalar; the techniques of Predictable Components enables one to transform an output vector Y to independent scalar variables $W_1, \ldots, W_d$.

III. Principal Components. The problem of modeling a d-dimensional random vector Y can perhaps be reduced to the problem of modeling a scalar random variable by transforming to d uncorrelated random variables $W_1, \ldots, W_d$ which are linear combinations of Y. Let $V_1, \ldots, V_d$ and $\lambda_1, \ldots, \lambda_d$ be the latent vectors and latent values of Γ_{YY} satisfying: $\Gamma_{YY} V_j = \lambda_j V_j$, $j = 1, \ldots, d$; $\lambda_1 > \lambda_2 > \ldots > \lambda_d$; and $V_j' V_j = 1$. Then $W_j = V_j' Y$ is called the j^{th} principal component of Y; it can be characterized as the linear combination of Y (whose coefficients have sum of squares equal to 1) which has maximum variance and is uncorrelated with W_k, $k < j$. Note that the variance of W_j is λ_j. A more useful technique for modeling seems to be canonical correlations and the components to which they lead, which we call predictable components.

IV. Predictable Components and Canonical Correlations. To gain insight into the regression representation

$$Y = Y^{\mu} + Y^{\nu}$$

where $Y^{\mu} = BX$ and Y^{ν} has covariance matrix Σ, one should transform Y to d uncorrelated random variables $W_1, \ldots, W_d$ defined as follows: W_1 is the most predictable linear combination of Y from X; W_2 is the next most predictable, and so on. Writing $W = m'Y$, the unpredictability measure of W is

$$\lambda = \frac{\operatorname{Var}[m'Y^{\nu}]}{\operatorname{Var}[m'Y]} = \frac{m'\Sigma m}{m'\Gamma_{YY} m}$$

It is shown in Rao (1973), p. 74 that the vector m which minimizes λ is the latent vector corresponding to the latent value λ of

$$\Gamma_{YY}^{-1}\Sigma = I - \Gamma_{YY}^{-1}\Gamma_{YX}\Gamma_{XX}^{-1}\Gamma_{XY} .$$

Equivalently m and λ satisfy

$$\Sigma m = \lambda \Gamma_{YY} m .$$

Let V_j be the latent vector satisfying

$$\Gamma_{YY}^{-1}\Gamma_{YX}\Gamma_{XX}^{-1}\Gamma_{XY} V_j = (1 - \lambda_j) V_j, \quad j = 1, \ldots, d$$

where $1 - \lambda_1 < 1 - \lambda_2 < \ldots < 1 - \lambda_d$. Then $V_j'Y$ has predictor $V_j'Y^{\mu} = V_j'BX = U_j'X$ where $U_j = B'V_j$ satisfies

$$\Gamma_{XX}^{-1}\Gamma_{XY}\Gamma_{YY}^{-1}\Gamma_{YX}U_j = (1-\lambda_j)\,U_j \;.$$

The prediction error $V_j'Y^{\nu}$ has variance

$$V_j'\Sigma V_j = \lambda_j V_j'\Gamma V_j \;.$$

Thus $1-\lambda_j$ can be interpreted as the square of a correlation coefficient, called a canonical correlation (see Brillinger (1975)).

We call $W_j = V_j'Y$ the j^{th} predictable component of Y .

V. Subset Regression of Predictable Components. Instead of predicting $W_j = V_j'Y$ by the linear combination $U_j'X$, where $U_j = B'V_j$, one seeks for each $j = 1,\ldots,d$ a parsimonious predictor of W_j given X , using subset regression techniques. One might also interpret some of the coefficients of V_j as being zero if they are small enough. In this way one finds relations

$$c_{11}Y_1 + \ldots + c_{1d}Y_d = b_{11}X_1 + \ldots + b_{1r}X_r + \epsilon_1$$

$$c_{21}Y_1 + \ldots + c_{2d}Y_d = b_{21}X_1 + \ldots + b_{2r}X_r + \epsilon_2$$

and so on, where each equation has as few non-zero coefficients as possible, the variances of $\epsilon_1, \epsilon_2, \ldots$ are successively larger, and $\epsilon_1, \ldots \epsilon_d$ are approximately uncorrelated.

3. Auto-Regression and Cross-Regression Representations

Modeling a multiple time series $Y(t)$ begins with representations as a sum

$$Y(t) = Y^{\mu}(t) + Y^{\nu}(t)$$

where $Y^{\mu}(t)$ is the predictor of $Y(t)$ given a certain information base, and $Y^{\nu}(t)$ is always defined by $Y^{\nu}(t) = Y(t) - Y^{\mu}(t)$.

We define $Y^{\mu}(t)$ to be the conditional expectation of $Y(t)$ given a set of explanatory or predictor variables which we denote PREDVAR:

$$Y^{\mu}(t) = E[Y(t)\,|\,\text{PREDVAR}] \;.$$

Instead of a regression matrix B such that $Y^{\mu} = B \cdot \text{PREDVAR}$, one must specify a linear operator or filter, called the PREDFIL or prediction filter, which transforms the PREDVAR series to $Y^{\mu}(\cdot)$. In symbols

$$Y^{\mu} = \text{PREDFIL} \cdot \text{PREDVAR}$$

or

$$\text{PREDVAR} - \boxed{\text{PREDFIL}} - Y^{\mu}(t) \;.$$

Instead of an error covariance matrix Σ, one specifies a sequence, called PREDSIGMA, of prediction error covariance matrices

$$\Sigma(v) = E[Y^{\nu}(t)\{Y^{\nu}(t+v)\}^*]$$

which describes the probability law of $Y^{\nu}(\cdot)$ when it is normal. We denote $\Sigma(0)$ by Σ. Often $\Sigma(v) = 0$ for $v \neq 0$; we then call $Y^{\nu}(\cdot)$ multiple white noise.

Parameter identification is concerned with the estimation of PREDFIL and PREDSIGMA *for a given choice of* PREDVAR. Model identification is concerned with choosing PREDVAR; as in subset regression, this is often done by suitable comparisons of PREDSIGMA corresponding to different choices of PREDVAR.

Some typical choices for PREDVAR (to predict $Y(t)$ at a fixed time t) are:

Case Auto (infinite): The predictor variables consist of the infinite past $Y(t-1)$, $Y(t-2), \ldots$: $Y^{\mu}(t) = -\sum_{j=1}^{\infty} A_{YY}(j)\, Y(t-j)$.

Case Auto (m): The predictor variables consist of the finite past $Y(t-1), \ldots, Y(t-m)$: $Y^{\mu}(t) = -\sum_{j=1}^{m} A_{YY}(j)\, Y(t-j)$.

Case Cross (infinite): The predictor variables consist of the infinite past and future values of "input" variables $X(s)$, $s = 0, \pm 1, \ldots$ which are also a normal, zero mean, stationary multiple time series:

$$Y^{\mu}(t) = \sum_{k=-\infty}^{\infty} A_{YX}(k)\, X(t-k) .$$

Case Cross (finite): The predictor variables consist of a finite set of past and future values of $X(s)$, $t - m_1 \leq s \leq t + m_2$ where m_1 and m_2 are fixed positive integers: $Y^{\mu}(t) = \sum_{k=-m_1}^{m_2} A_{YX}(k)\, X(t-k)$.

Case Auto-cross (infinite past): The predictor variables consist of the infinite past $Y(t-1)$, $Y(t-2) \ldots$ of the "output" variables, and the infinite past $X(t-1)$, $X(t-2), \ldots$ of the "input" variables: $Y^{\mu}(t) = -\sum_{j=1}^{\infty} A_{YY}(j)\, Y(t-j) + \sum_{k=1}^{\infty} A_{YX}(k)\, X(t-k)$.

Case Auto-cross (m): The predictor variables consist of $Y(t-1), \ldots, Y(t-m)$, $X(t-1), \ldots, X(t-m)$: $Y^{\mu}(t) = -\sum_{j=1}^{m} A_{YY}(j)\, Y(t-j) + \sum_{k=1}^{m} A_{YX}(k)\, X(t-k)$.

Case Auto-cross-cross (m): The predictor variables add $X(t)$ to $Y(t-1), \ldots, Y(t-m)$, $X(t-1), \ldots, X(t-m)$; $Y^{\mu}(t) = -\sum_{j=1}^{m} A_{YY}(y)\, Y(t-j) + \sum_{k=0}^{m} A_{YX}(k)\, X(t-k)$.

Formulas for PREDFIL and PREDSIGMA are discussed in the sequel. To identify the foregoing models, one basic approach is to model a multiple time series $Z(t)$ defined as follows: when $Y(\cdot)$ is modeled using its own past, $Z(t) = Y(t)$; when modeling $Y(\cdot)$ also using the past of $X(\cdot)$,

$$Z(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} .$$

4. The Spectral Method

The spectral density matrix $f(\omega)$, $-\pi \le \omega \le \pi$, of a stationary multiple time series $Y(\cdot)$ with covariance matrix function $R(v)$ is defined by

$$f(\omega) = \frac{1}{2\pi} \sum_{v=-\infty}^{\infty} e^{-iv\omega} R(v)$$

which leads to the representation

$$R(v) = \int_{-\pi}^{\pi} e^{iv\omega} f(\omega)\, d\omega \quad .$$

An estimator of $f(\omega)$ is denoted $\hat{f}(\omega)$.

For a fixed frequency ω , $f(\omega)$ is a covariance matrix and $\hat{f}(\omega)$ is a sample covariance, whose distribution can usually be assumed to be approximately a complex Wishart matrix (a review of complex multivariate distributions and their applications in some problems of inference on multiple time series is given by Krishnaiah (1976)).

It is appealing to carry out the techniques of conventional multivariate analysis on $\hat{f}(\omega)$ at each frequency ω; the theory for this is given by Brillinger (1975). From a practical point of view one would like to interpret in the time domain the analysis carried out in the spectral domain.

Spectral methods are appropriate for estimating PREDVAR and PREDSIGMA in the Cross case, in which $Y^{\mu}(\cdot)$ is only a function of a series $X(\cdot)$. Then one estimates the spectral density matrix

$$f_Z(\omega) = \begin{bmatrix} f_{XX}(\omega) & f_{XY}(\omega) \\ f_{YX}(\omega) & f_{YY}(\omega) \end{bmatrix}$$

Define the filter transfer function

$$B(\omega) = \sum_{k=-\infty}^{\infty} e^{-i\omega k} A_{YX}(k)$$

One can show that

$$B(\omega) = f_{YX}(\omega)\, f_{XX}^{-1}(\omega)$$

By matrix pivoting techniques one can transform $f_Z(\omega)$ to

$$\begin{bmatrix} f_{XX}^{-1}(\omega) & \overline{B}(\omega) \\ -B(\omega) & f_{\epsilon\epsilon}(\omega) \end{bmatrix}$$

where

$$f_{\epsilon\epsilon}(\omega) = f_{YY}(\omega) - f_{YX}(\omega)\, f_{XX}^{-1}(\omega)\, f_{XY}(\omega) \quad ,$$

Finally $A_{YX}(j)$ and Σ are found by

$$A_{YX}(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} B(\omega)\, e^{ij\omega}\, d\omega \quad ,$$

$$\Sigma = \int_{-\pi}^{\pi} f_{\epsilon\epsilon}(\omega)\, d\omega \quad .$$

When estimating $A_{YX}(j)$ from an estimator of $B(\omega)$ one can use a regression approach (see Parzen (1967a) and equation (2) below).

To estimate PREDFIL and PREDSIGMA in the Cross case, one needs to first estimate $f_Z(\omega)$ which can be done by two methods: windowed covariances and autoregressive. We prefer the latter and describe it in the next section. The remainder of this section discusses some difficulties in implementing in the frequency domain the "principal components" method of conventional multivariate analysis.

Index Time Series. The multiple time series analyst aims to display the information in a d-dimensional time series by summarizing it in a series of reduced dimension. In particular one seeks to find a univariate time series which in some sense best summarizes the information in a d-dimensional series. One approach to this problem is through principal components in the frequency domain.

One method of defining an index series is to form a scalar series

$$W(t) = \sum_{j=-\infty}^{\infty} b^T(j)\, Y(t-j)$$

and an approximating series to $Y(t)$ denoted

$$Y^a(t) = \sum_{j=-\infty}^{\infty} c(j)\, W(t-j) \quad ,$$

where the coefficient matrices $b(\cdot)$ and $c(\cdot)$ are determined to minimize the mean square approximation error $E(Y(t) - Y^a(t))^*(Y(t) - Y^a(t))$ Brillinger (1975) shows that the $b(\cdot)$ and $c(\cdot)$ are given by

$$b(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} V_1(\omega)\, e^{ij\omega}\, d\omega = c(j) \qquad (1)$$

where $V_1(\omega)$ is the eigenvector corresponding to the largest eigenvalue of the spectral density matrix $f_{YY}(\omega)$ of Y. Further, the minimum mean square approximation error is equal to

$$\int_{-\pi}^{\pi} \sum_{j=2}^{d} \lambda_j(\omega)\, d\omega \quad .$$

18

where $\lambda_1(\omega), \ldots, \lambda_d(\omega)$ are the eigenvalues of $f_{YY}(\omega)$.

To estimate the $b(\cdot)$, Brillinger suggests using a nonparametric estimator of $f_{YY}(\cdot)$, finding its largest eigenvector, and approximating the integral in (1) by rectangular sums (which can be done via the fast Fourier transform algorithm). We show how to use a regression approach to estimate $b(j)$ from an estimator of $V_1(\omega)$. However it involves a complex valued regression problem with heteroskedastic errors having singular covariance matrices. Further $V_1(\omega)$ is only determined up to a complex-valued factor of modulus 1 .

<u>Theorem</u> (Brillinger (1975), p. 351)

Let $f_{YY}^{(T)}(\omega)$ be a nonparametric estimator of $f_{YY}(\omega)$ with weighting function $K(\cdot)$, and let B_T be a nonnegative bandwidth parameter. Let $\lambda_1^{(T)}(\omega), \ldots, \lambda_d^{(T)}(\omega)$ and $\underset{\sim}{V}_1^{(T)}(\omega), \ldots, \underset{\sim}{V}_d^{(T)}(\omega)$ be the eigenvalues and eigenvectors of $f_{YY}^{(T)}(\omega)$. Then

$$\lim_{T\to\infty} B_T T \operatorname{Cov}(V_1^{(T)}(\omega_j), V_1^{(T)}(\omega_k))$$

$$= \delta_{jk} 2\pi \int K^2(\omega)\, d\omega\, \lambda_1(\omega) \sum_{\ell=2}^{d} \frac{\lambda_\ell(\omega)}{[\lambda_1(\omega) - \lambda_\ell(\omega)]^2} V_1(\omega) V_1^*(\omega)$$

$$= C(\omega) ,$$

where δ_{jk} is the Kronecker delta and $\omega = \omega_j = \omega_k$. Thus this limiting $(d \times d)$ covariance matrix is of rank at most $d-1$ and is thus singular. Further, it is complex valued.

Since $V_1(\cdot)$ and $b(\cdot)$ are Fourier pairs, we can write

$$V_1^{(T)}(\omega) = V_1(\omega) + \epsilon(\omega)$$

where

$$V_1(\omega) = \sum_{j=-\infty}^{\infty} b(j) e^{-ij\omega} ,$$

and

$$\operatorname{Cov}(\epsilon(\omega_j), \epsilon(\omega_k)) \to \delta_{jk} C(\omega) .$$

Thus one could determine estimators of the $b(\cdot)$ by a regression approach <u>i.e.</u> minimize the errors in

$$V_1^{(T)}(\omega_k) = \sum_{j=-M}^{M} b_{T,M}(j) e^{-ij\omega_k} + \epsilon_{T,M}(\omega_k) , \qquad (2)$$

$k = 1, \ldots, Q$, giving rise to the equation

$$y = X\beta + \epsilon$$

where $y = (V_1^{(T)}(\omega_1), \ldots, V_1^{(T)}(\omega_Q))^T$, the $(j,k)^{th}$ block element of X is $X_{jk} = e^{i(M-k+1)\omega_j} I_d$, $j = 1, \ldots, Q, k=1, \ldots, 2M+1$, $\beta = (b_{T,M}^T(-M), \ldots, b_{T,M}^T(M))^T$, and Cov ϵ is asymptotically the block diagonal matrix C having $C(\omega_k)$ as the k^{th} block diagonal element. Thus one could find the complex valued weighted least squares estimator of β using a generalized inverse of C . But this would lead to complex valued estimators of the $b(\cdot)$. Note that if one assumes $C(\omega_k) = I_d$, then one obtains the estimators given in Brillinger.

In an attempt to obtain real valued estimators of the $b(\cdot)$, one could perform a regression on the real and imaginary parts of $V_1^{(T)}(\omega)$. This would require knowing the covariance matrix of the vector of real and imaginary parts of the complex errors. That is we have a complex random variable $Z = X + iY$ where $E(ZZ^*) = \Sigma$ and we want Cov $\begin{bmatrix} X \\ Y \end{bmatrix}$. If one adds the assumption that $E(ZZ') = 0$, then

$$\mathrm{Cov}\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{2}\begin{bmatrix} \mathrm{Re}\,\Sigma & -\mathrm{Im}\,\Sigma \\ \mathrm{Im}\,\Sigma & \mathrm{Re}\,\Sigma \end{bmatrix}$$

With this condition, the normal equations can be solved using a generalized inverse of the covariance matrices.

However, all of the above methods lead to nonunique estimators since they are of the form (where $w(\cdot)$ are weighting matrices)

$$\hat{b}(j) = \sum_k w(k) V_1^{(T)}(\omega_k)$$

and if $V_1^{(T)}(\omega_k)$ is a unit length eigenvector of $f_{YY}^{(T)}(\omega_k)$, then so is $e^{i\theta} V_1^{(T)}(\omega_k)$ for any θ .

Bloomfield (1978) proposes a procedure for making a unique identification of $V_1(\omega)$. However it is still difficult to interpret the index series $W(t)$ formed from the filter whose frequency transfer function is $V_1(\omega)$.

In the worked example in Brillinger (1975), p. 355, the multiple time series consists of components which have been pre-whitened by the removal of seasonal effects. The spectral density matrix $f_{YY}(\omega)$ is then real, and $b(j) = 0$ for $j \neq 0$. The index series $W(t)$ is a linear function only of the values of $Y(t)$ at the same time t , and indeed turns out to be their average.

5. <u>The Autoregressive Method</u>

Given a sample realization $Z(1), \ldots, Z(T)$ from a d-dimensional time series $Z(\cdot)$, the p^{th} order autoregressive approximating spectral estimator is defined (see Parzen (1969)):

$$\hat{f}_p(\omega) = \frac{1}{2\pi} G_p^{-1}(e^{i\omega}) \Sigma_p G_p^{-*}(e^{i\omega}) \quad , \quad \omega \in [-\pi, \pi]$$

where $G_p(z)$ is the complex matrix polynomial

$$G_p(z) = \sum_{j=0}^{p} A_p(j) z^j \quad ;$$

$A_p(0) = I$; and $A_p(1), \ldots, A_p(p)$ and Σ_p are found from the sample Yule-Walker equations

$$\sum_{j=0}^{p} A_p(j) R_T(j-v) = \delta_{v,0} \Sigma_p \quad , \quad v = 1, \ldots, p \quad .$$

$R_T(v)$ is the sample covariance function defined by

$$R_T(v) = \frac{1}{T} \sum_{t=1}^{T-v} Z(t) Z^*(t+v) = R_T^*(-v) \quad .$$

An optimal order $\hat{p}$ is determined as the value of p which minimizes an order determining criterion such as Akaike's AIC criterion (see Akaike (1976)) or

$$CAT(p) = tr\left[\frac{d}{T} \sum_{j=1}^{p} \frac{T-jd}{T} \Sigma_j^{-1} - \frac{T-pd}{T} \Sigma_p^{-1}\right] \quad ,$$

where $tr(A)$ is the trace of the matrix A, introduced in Parzen (1977).

To make inferences about PREDFIL and PREDSIGMA, one uses mainly the time domain parameters $\hat{p}$, $A_{\hat{p}}(\cdot)$, $\Sigma_{\hat{p}}$, $R_T(\cdot)$; the use of the frequency domain parameters $f_{\hat{p}}(\omega)$ and spectral parameters derived from $f_{\hat{p}}(\omega)$ was illustrated in the previous section. We now summarize the autoregressive method of modeling Y using the representation $Y = Y^{\mu} + Y^{\nu}$.

The <u>auto</u> case assumes $Y^{\mu}(t)$ is only a function of past $Y(\cdot)$; then its parameters can be obtained by applying the autoregressive method to $\cdot$ $Y(\cdot)$.

The <u>cross</u> case assumes $Y^{\mu}(t)$ is only a function of a series $X(\cdot)$ then estimators are obtained by applying the autoregressive method to form an estimator $\hat{f}_Z(\omega)$ from which one forms $\hat{B}(\omega)$ and $\hat{f}_{\epsilon\epsilon}(\omega)$ which are used to estimate PREDFIL and PREDSIGMA (the formulas are given in the preceding section).

The <u>auto-cross</u> case assumes $Y^{\mu}(t)$ is a function of past $Y(\cdot)$ and past $X(\cdot)$; then estimators $\hat{p}$, $A_{YY,\hat{p}}(j)$, $A_{YX,\hat{p}}(j)$ and $\Sigma_{YY,\hat{p}}$ can be found by fitting an autoregressive approximant to $Z(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}$;

$$\sum_{j=0}^{\hat{p}} A_{\hat{p}}(j)\, Z(t-j) = \epsilon(t)$$

where $\epsilon(t) = (\epsilon_X^*(t), \epsilon_Y^*(t))^*$, and

$$A_{\hat{p}}(j) = \begin{bmatrix} A_{XX,\hat{p}}(j) & A_{XY,\hat{p}}(j) \\ A_{YX,\hat{p}}(j) & A_{YY,\hat{p}}(j) \end{bmatrix} .$$

$$\Sigma_{\epsilon\epsilon,\hat{p}} = \begin{bmatrix} \Sigma_{XX,\hat{p}} & \Sigma_{XY,\hat{p}} \\ \Sigma_{YX,\hat{p}} & \Sigma_{YY,\hat{p}} \end{bmatrix} .$$

Thus $Y(t) = Y^{\mu}(t) + Y^{\nu}(t)$ where

$$Y^{\mu}(t) = -\sum_{j=1}^{\hat{p}} A_{YY,\hat{p}}(j)\, Y(t-j) + \sum_{j=1}^{\hat{p}} A_{YX,\hat{p}}(j)\, X(t-j) , \qquad (1)$$

and $\mathrm{Var}\, Y^{\nu}(t) = \Sigma_{YY,\hat{p}}$. Similarly one can form $X^{\mu}(t)$ and $X^{\nu}(t) = X(t) - X^{\mu}(t)$.

The <u>auto-cross-cross</u> case adds $X(t)$ to the predictor set for $Y(t)$. Parzen (1969) shows that one can do this by forming a predictor of $Y^{\nu}(t)$, denoted by $Y^{\nu+}(t)$:

$$Y^{\nu+}(t) = \Sigma_{YX,\hat{p}} \Sigma_{XX,\hat{p}}^{-1} X^{\nu}(t)$$

where $X^{\nu}(t) = X(t)$ - predictor of $X(t)$ given past X and Y. The predictor of $Y(t)$ given the past of Y , and the past and present of X , is

$$Y^{\mu+}(t) = Y^{\mu}(t) + Y^{\nu+}(t)$$

where $Y^{\mu}(t)$ is given by (0) .

The estimated covariance matrix of $Y(t) - Y^{\mu+}(t)$ is given by

$$\Sigma_{YY,\hat{p}} - \Sigma_{YX,\hat{p}} \Sigma_{XX,\hat{p}}^{-1} \Sigma_{XY,\hat{p}} .$$

It is important to note that one can vary which components of Z belong to Y and which to X quite simply by using the autoregressive method.

As a simple illustration of how joint autoregressive modeling provides an approach to fitting and estimating relations between time series, consider two scalar series $Y(t)$ and $X(t)$ satisfying

$$\begin{aligned} Y(t) &= \gamma_0 X(t) + \gamma_1 X(t-1) + \eta(t) \\ X(t) - \rho X(t-1) &= \delta(t) \end{aligned} \qquad (2)$$

where $\delta(\cdot)$ and $\eta(\cdot)$ are independent white noise processes. Then

$$Z(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \quad \text{obeys}$$

$$A(0)\,Z(t) + A(1)\,Z(t-1) = \begin{bmatrix} \delta(t) \\ \eta(t) \end{bmatrix}$$

where

$$A(0) = \begin{bmatrix} 1 & 0 \\ -\gamma_0 & 1 \end{bmatrix}, \quad A(1) = \begin{bmatrix} -\rho & 0 \\ -\gamma_1 & 0 \end{bmatrix}$$

One may write

$$Z(t) + A\,Z(t-1) = \epsilon(t), \quad \text{Var}\ \epsilon = \Sigma ,$$

where

$$A = -\begin{bmatrix} \rho & 0 \\ \rho\gamma_0 + \gamma_1 & 0 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} \sigma_\delta^2 & \sigma_\delta^2\gamma_0 \\ \gamma_0\sigma_\delta^2 & \gamma_0^2\sigma_\delta^2 + \sigma_\eta^2 \end{bmatrix}$$

If $Z(t)$ obeys an autoregressive scheme of order 1, and its coefficient matrix A has zeroes in its second column, then one can represent the time series by the regression model in equation (2).

6. Predictable Components

To model and analyze a univariate time series Y, one approach is the method of decomposition, which attempts to decompose the time series as a *sum of terms (representing trend, seasonal, and irregular* as an example). To model and analyze a d-dimensional multiple time series $Y(t)$, one can adopt an analogous approach which writes Y as a linear transformation of a d-dimensional vector

$$W(t) = \begin{bmatrix} W_1(t) \\ \cdots \\ W_d(t) \end{bmatrix}$$

whose components are uncorrelated in the sense that its zero lag covariance matrix is a diagonal matrix D:

$$R_W(0) = E[\,W(t)\,W^*(t)\,] = D .$$

One chooses the components $W_1(t), \ldots, W_d(t)$ on the basis of predictability using a set PREDVAR of prediction variables. One chooses: $W_1(t)$ to be the most predictable linear combination of $Y(t)$; $W_2(t)$ to be the most predictable linear combination of $Y(t)$ which is uncorrelated with $W_1(t)$; $W_3(t)$ to be the most predictable linear combination of $Y(t)$ which is uncorrelated with $W_1(t)$ and $W_2(t)$; and so on.

For any linear combination $c'Y(t)$, its unpredictability criterion is

$$\lambda = \frac{\mathrm{Var}[c'Y^{\nu}(t)]}{\mathrm{Var}[c'Y(t)]} = \frac{c'\Sigma c}{c'R_Y(0)c}$$

The coefficient vector c with smallest value of λ is the latent vector corresponding to the <u>smallest</u> latent value λ_1 of $R_Y^{-1}(0)\Sigma$. Further λ_1 equals the unpredictability criterion λ of the most predictable linear combination of $Y(t)$.

For a specified set PREDVAR of prediction variables, one needs only Σ and $R_Y(0)$ to form the coefficient vectors needed to form the predictable components

$$W_j(t) = c_j'Y(t) \quad , \quad j = 1,\ldots,d \; ,$$

where $c_1,\ldots,c_d$ are the latent vectors corresponding to the ordered latent values $\lambda_1 < \lambda_2 < \ldots < \lambda_d$ of $R_Y^{-1}(0)\Sigma$:

$$R_Y^{-1}(0)\Sigma c_j = \lambda_j c_j \quad .$$

When λ_1 is close to 0 (say, smaller than 8/T) then we consider $W_1(t)$ to be a <u>highly smooth</u> time series and therefore an <u>index</u> time series. When λ_d is close to 1 (say, $1 - \lambda_d$ smaller than 8/T) we consider $W_d(t)$ to be a <u>white noise</u> series. For further details on the interpretations to be placed on the predictable components, see Box and Tiao (1977). The ideas of this section are an extension of their ideas.

Comparing the graphs of the time series $Y_1(t),\ldots,Y_d(t)$ with $W_1(t),\ldots,W_d(t)$ is very helpful in clarifying how much $W_1(t)$ summarizes the common smooth behavior of the original time series.

In summary, given a set PREDVAR of prediction variables for a multiple time series $Y(t)$, one can form the prediction filter PREDFIL , the prediction mean square error matrix PREDSIGMA , and the predictable components $W_1(t),\ldots,W_d(t)$. They may be interpretable. Further by comparing these estimated parameters for different choices of PREDVAR , one may be able to determine which sets of "input" (prediction) variables are most informative about the "output" variables $Y(t)$.

The present paper does not examine in detail the prediction filters, even though these should be estimated for a complete model identification; rather it emphasizes the insight to be obtained by determining the index time series $W_1(t)$. An illustrative example has been worked out; it involves monthly ozone level time series discussed in Parzen and Pagano (1978).

30

REFERENCES

Akaike, H. (1976). "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion," in System Identification, edited by R. K. Mehra and D. G. Mainiotis, New York: Academic Press, pp. 27-96.

Bloomfield, P. (1978). "Principal Components of Time Series," presented at Ames (Iowa) Special Meeting on Time Series Analysis, organized by the Institute of Mathematical Statistics.

Box, G. E. P. and Tiao, G. C. (1977). "A Canonical Analysis of Multiple Time Series," Biometrika, 64, 355-365.

Brillinger, D. R. (1975). Time Series Data Analysis and Theory. New York: Holt, Rinehart and Winston.

Krishnaiah, P. R. (1976). Some recent developments on complex multivariate distributions. Journal of Multivariate Analysis, Vol. 6, pp. 1-30.

Mehra, R. K. and Lainiotis, D. G. (1976). System Identification: Advances and Case Studies. New York: Academic Press.

Parzen, E. (1967). "Empirical Multiple Time Series Analysis," Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, edited by L. LeCam and J. Neyman, University of California Press, Vol. I, pp. 305-340.

__________ (1967a). "Time Series Analysis for Models of Signal Plus White Noise," in Spectral Analysis of Time Series, edited by B. Harris, New York: Wiley, 233-257.

__________ (1969). "Multiple Time Series Modeling," Multivariate Analysis - II, edited by P. Krishnaiah, New York: Academic Press, 389-409.

__________ (1977). "Multiple Time Series: Determining the Order of Approximating Autoregressive Schemes," Multivariate Analysis - IV, edited by P. Krishnaiah, New York: North Holland Publishing, 283-295.

Parzen, E. and Pagano, M. (1978). "Statistical Time Series Analysis of Worldwide Total Ozone for Trends." Unpublished manuscript.

Priestley, M. B. (1976). "Applications of Multivariate Techniques in the Study of Multivariate Stochastic Systems." Paper presented at the Sixth International Conference on Stochastic Processes at Tel Aviv, Israel.

__________ (1968). "System Identification, Kalman Filtering, and Stochastic Control," presented at Ames (Iowa), Special Meeting on Time Series Analysis organized by the Institute of Mathematical Statistics.

Priestley, M. B., Subba Rao, T., and Tong, H. (1973). "Identification of the Structure of Multivariate Stochastic Systems," Multivariate Analysis III (ed. by P. R. Krishnaiah). New York: Academic Press, 351-368.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. New York: Wiley.

Subba Rao, T. (1975). "An Innovation Approach to the Reduction of the Dimensions in a Multivariate Stochastic System," International Journal of Control, 673-680.

APPENDIX

Example: Monthly Ozone Levels

The Data

In this section we present the results of analyzing monthly total ozone levels at nine observation stations throughout the world (see Hill, Tiede, Sheldon (1977)). Plots of the original data are given in the figures.

Prewhitening

The number of observations making up a monthly total level is a function of how clear the skies are above a particular station. Since this could be a function of time of year, prewhitening by merely subtracting an overall mean may be insufficient.

Thus we have examined the results for each station by applying the following detrending and prewhitening options (with parameters calculated with the last 24 data points excluded for each series):

(1) Subtract overall mean.

(2) Subtract monthly means and divide by monthly standard deviations.

(3) Subtract monthly fitted means and divide by monthly standard deviations.

(4) Apply autoregressive filter to monthly mean and variance detrended series, with AR order determined by CAT.

(5) Apply autoregressive filter to monthly fitted mean and variance detrended series, with AR order determined by CAT.

Multiple Analysis

Three of the stations (Huancayo, Kodaikanal, and Mauna Loa) are near the equator and thus there is less seasonal variability than in the six nonequatorial stations (Arosa, Aspendale, Buenos Aires, Edmonton, Macquarie Islands, and Tateno). Thus for each method of prewhitening we have considered analyses of $Z(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}$ where Y is the series of nonequatorial stations and X is the series of equatorial stations; one can consider five cases:

(1) $Z(t)$ AR fitted by its past, denoted $(Z|Z)$

(2) $Y(t)$ AR fitted by its past, denoted $(Y|Y)$

(3) $Y(t)$ fitted by past and future of X, denoted $(Y|X)$

(4) $Y(t)$ fitted by past of Y and X, denoted $(Y|Y,X)$

(5) $Y(t)$ fitted by past of Y and past and present of X, denoted $(Y|Y,X^+)$

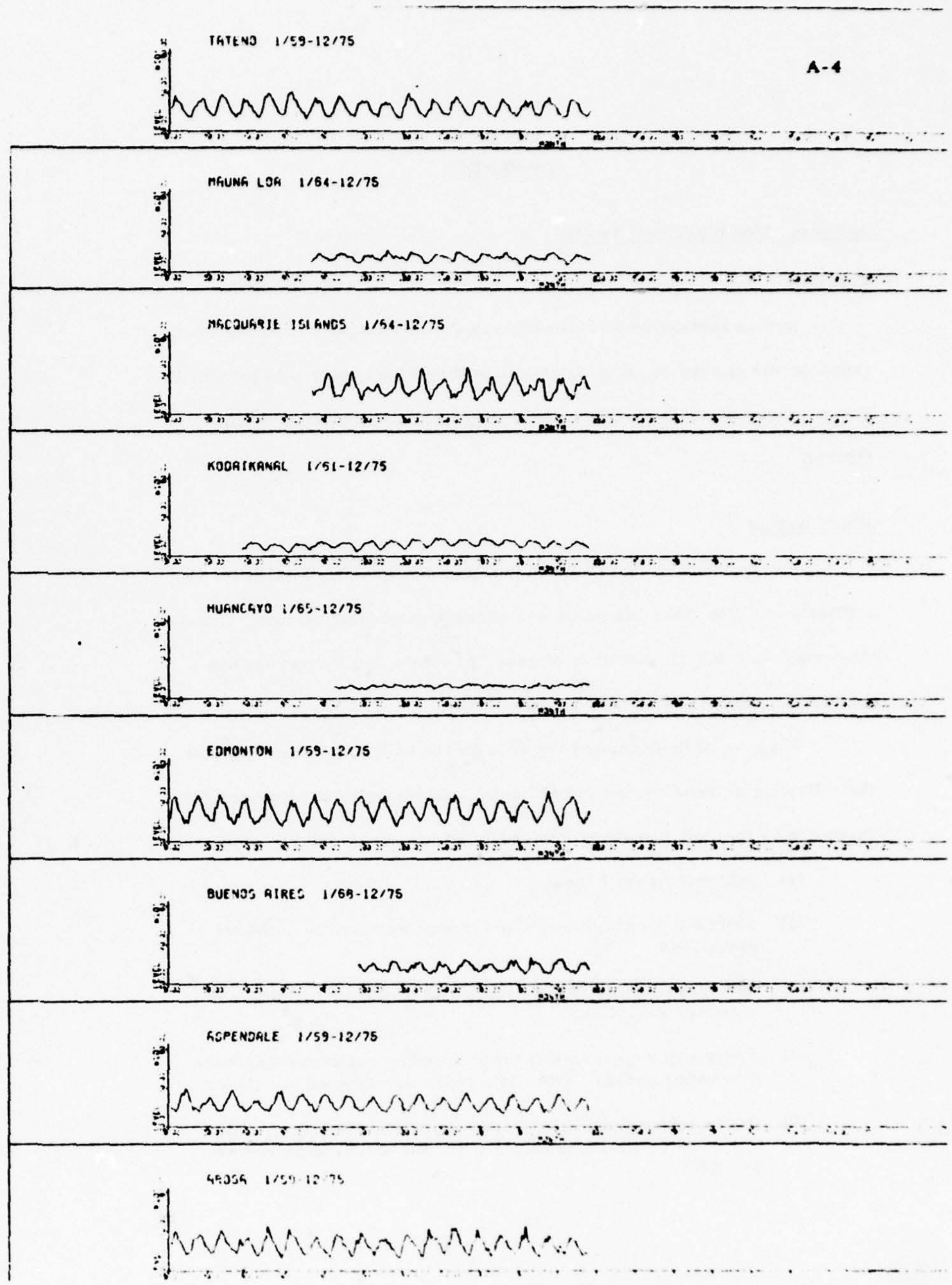
When one only substracts an overall mean, the series still contain periodic components with period 12. These are considered to be predictable components. Consequently, it is not surprising that the minimum unpredictability criterion λ is close to zero (more precisely λ equals .10 for $Y|Y$, .07 for $Y|Y,X$, .06 for $Y|Y,X^+$, and .08 for $Y|X$). The latent vector corresponding to the minimum latent value λ are similar for the three cases; it assigns weights close to 1 for two northern stations, Arosa and Tateno, and weight close to -1 for a southern station, Aspendale. Since southern stations have their minima

when northern stations have their maxima (180° out of phase), the minus sign puts the three stations in phase, and the most predictable component $W_1(t)$ exhibits a graph much like a typical northern station. However it is somewhat more predictable; indeed, the stations with the most weight in forming $W_1(t)$ are the most predictable stations.

When the time series are detrended (monthly means subtracted), the time series become only mildly auto-correlated and hardly cross-correlated. The predictable components W_3, W_4, W_5, and W_6 have unpredictability criteria greater than 0.8 and are considered white noise. The most predictable component W_1, which has unpredictability criterion of about 0.5 has coefficients close to zero for all stations except one (Kodaikanal) which happens to be a far more predictable station than other stations.

Reference

Hill, W. J., Sheldon, P. N. and Tiede, J. J. (1977), "Analyzing Worldwide Total Ozone for Trends," Geophysical Research Letters, 4, 21-24.



OVERALL MEAN ADJUSTMENT

$R_Z(0)$

Arosa	1.03								
Aspendale	-.72	.89							
Buenos Aires	-.57	.72	1.00						
Edmonton	.75	-.80	-.68	.92					
Macquarie Islands	-.72	.73	.60	-.72	.99				
Tateno	.74	-.67	-.50	.66	-.61	.81			
Huancayo	-.46	.41	.23	-.45	.32	-.48	.97		
Kodaikanal	.04	.24	.36	-.32	.23	-.05	.09	1.01	
Mauna Loa	.28	-.06	.03	-.05	-.06	.27	-.12	.67	1.06

Order of AR model fitted to Z is 1 .

$\Sigma_{Z|Z}$

Arosa	.18								
Aspendale	-.01	.14							
Buenos Aires	-.05	.08	.35						
Edmonton	.06	-.07	-.07	.18					
Macquarie Islands	-.02	.03	.01	-.05	.25				
Tateno	.01	-.03	-.03	.04	-.01	.14			
Huancayo	.01	.05	.04	-.06	-.02	-.02	.45		
Kodaikanal	-.02	.05	.03	-.05	.03	-.05	.11	.15	
Mauna Loa	-.02	.01	-.06	-.05	.02	-.04	.11	.07	.30

$\Sigma_{Y|Y}$

Arosa	.21					
Aspendale	-.04	.20				
Buenos Aires	-.08	.16	.45			
Edmonton	.09	-.15	-.16	.29		
Macquarie Islands	-.03	.07	.05	-.10	.28	
Tateno	.03	-.07	-.08	.10	-.04	.18

$\Sigma_{Y|X}$

Arosa	.22					
Aspendale	-.02	.13				
Buenos Aires	-.02	.05	.31			
Edmonton	.07	-.06	-.03	.17		
Macquarie Islands	-.08	.08	.04	-.07	.37	
Tateno	.03	-.03	.01	.03	-.03	.15

$\Sigma_{Y|Y,X}$ = 6th principal minor of $\Sigma_{Z|Z}$

$\Sigma_{Y|Y,X^+}$

Arosa	.18					
Aspendale	-.01	.12				
Buenos Aires	-.05	.06	.32			
Edmonton	.05	-.06	-.06	.16		
Macquarie Islands	-.01	.02	.01	-.04	.24	
Tateno	.01	-.02	-.02	.02	.00	.12

Latent Values of $R_Y^{-1}(0)\Sigma$

Z'Z	Y|Y	Y|X	Y|Y, X	Y|Y, X+
.04	.10	.08	.07	.06
.12	.46	.44	.39	.37
.37	.80	.79	.68	.64
.63	.83	.85	.79	.75
.71	.90	.90	.84	.82
.78	.99	.92	.90	.86
.90				
.92				
.94				

Latent Vectors (column vector) corresponding to increasing latent values

$\Sigma_{Z'Z}$

Arosa	.57	-.17	-.10	-.70	-.36	1.00	-.11	-.70	-.24
Aspendale	-.51	.43	-.69	-.73	-.03	.75	1.00	.53	-.09
Buenos Aires	.20	.25	-.37	-.32	-.02	.26	-.63	.24	.11
Edmonton	.01	-.20	-.11	-.49	1.00	.32	.10	1.00	-.56
Macquarie Islands	-.33	.24	-.09	-.05	.93	.15	-.16	-.52	-.06
Tateno	.82	.00	-.39	.32	.40	.10	.33	.05	1.00
Huancayo	-.41	-.31	.28	.76	.17	.70	-.10	.09	.14
Kodaikanal	1.00	1.00	1.00	-.45	.13	-.25	.18	.20	.13
Mauna Loa	.40	.20	-.72	1.00	.01	.03	-.09	.10	-.41

$\Sigma_{Y|Y}$

Arosa	1.00	.66	-.12	.51	.28	-.98
Aspendale	-.78	.53	-.21	.89	1.00	1.00
Buenos Aires	.19	.61	.13	.34	-.73	.00
Edmonton	-.47	-.21	1.00	1.00	.09	.75
Macquarie Islands	-.45	.59	.80	-.45	.13	-.45
Tateno	.83	1.00	-.11	-.84	.19	.96

$\Sigma_{Y|X}$

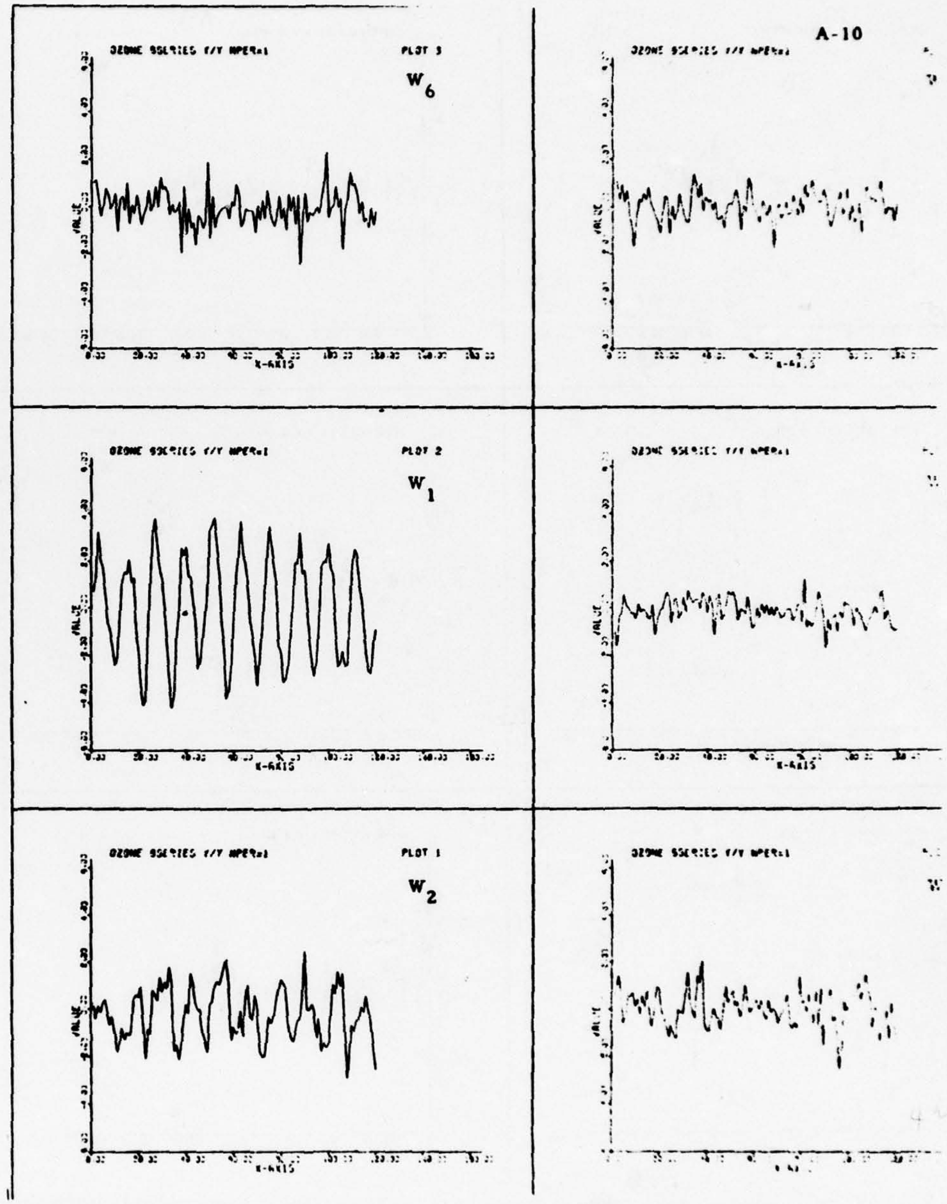
Arosa	-.59	1.00	-.22	.78	-.64	.20
Aspendale	1.00	.42	.51	-.58	.03	1.00
Buenos Aires	.39	.56	.43	.19	-.11	-.50
Edmonton	-.40	-.91	1.00	.42	.22	.29
Macquarie Islands	-.03	.05	-.14	1.00	.62	.00
Tateno	-.94	.92	-.01	-.79	1.00	.10

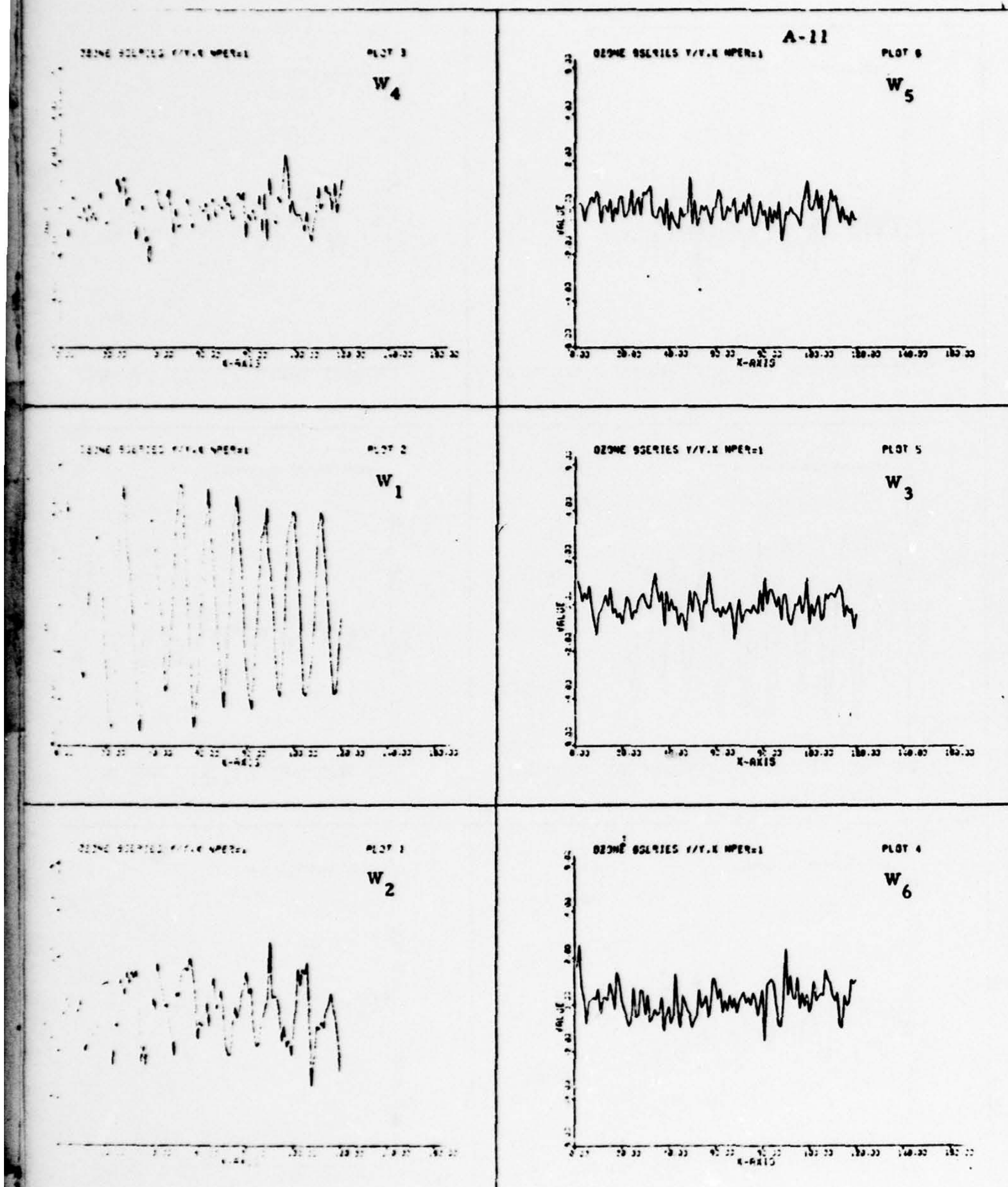
$\Sigma_{Y|Y,X}$

Arosa	.95	.93	-.31	-.67	-.46	-.58
Aspendale	-.97	.48	.39	-.72	1.00	-.99
Buenos Aires	.02	.71	.05	-.25	-.16	1.00
Edmonton	.01	-.65	1.00	-.94	.27	.26
Macquarie Islands	-.52	.30	.58	.25	-.60	-.28
Tateno	1.00	1.00	.37	1.00	.55	-.11

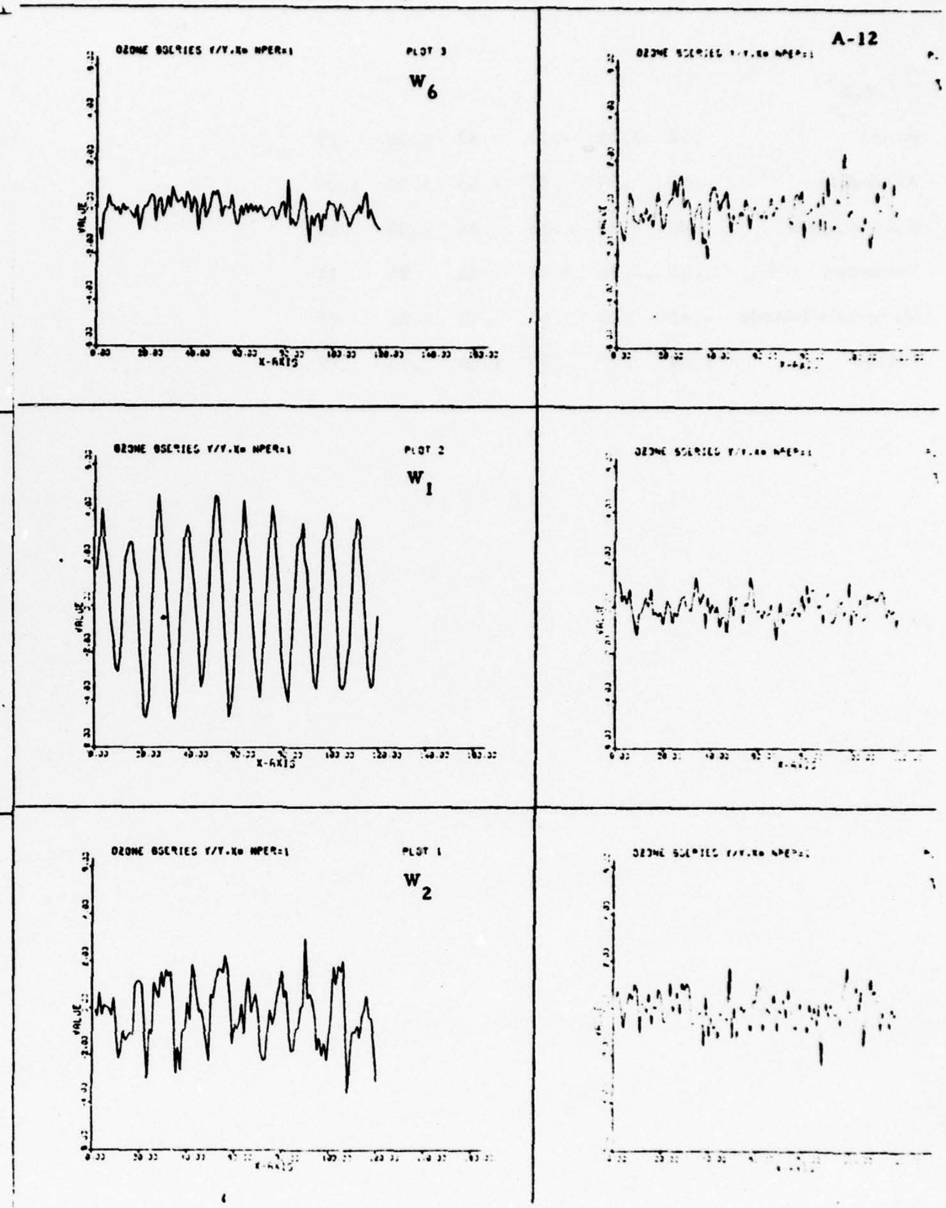
$\Sigma_{Y|Y,X^+}$

Arosa	.72	1.00	-.33	-.47	-.89	.29
Aspendale	-.89	.59	.27	-.65	1.00	1.00
Buenos Aires	.06	.84	.16	-.36	-.03	-.63
Edmonton	.12	-.56	1.00	-.98	.29	.11
Macquarie Islands	-.46	.20	.53	.45	-.86	.06
Tateno	1.00	.97	.30	1.00	.85	.17

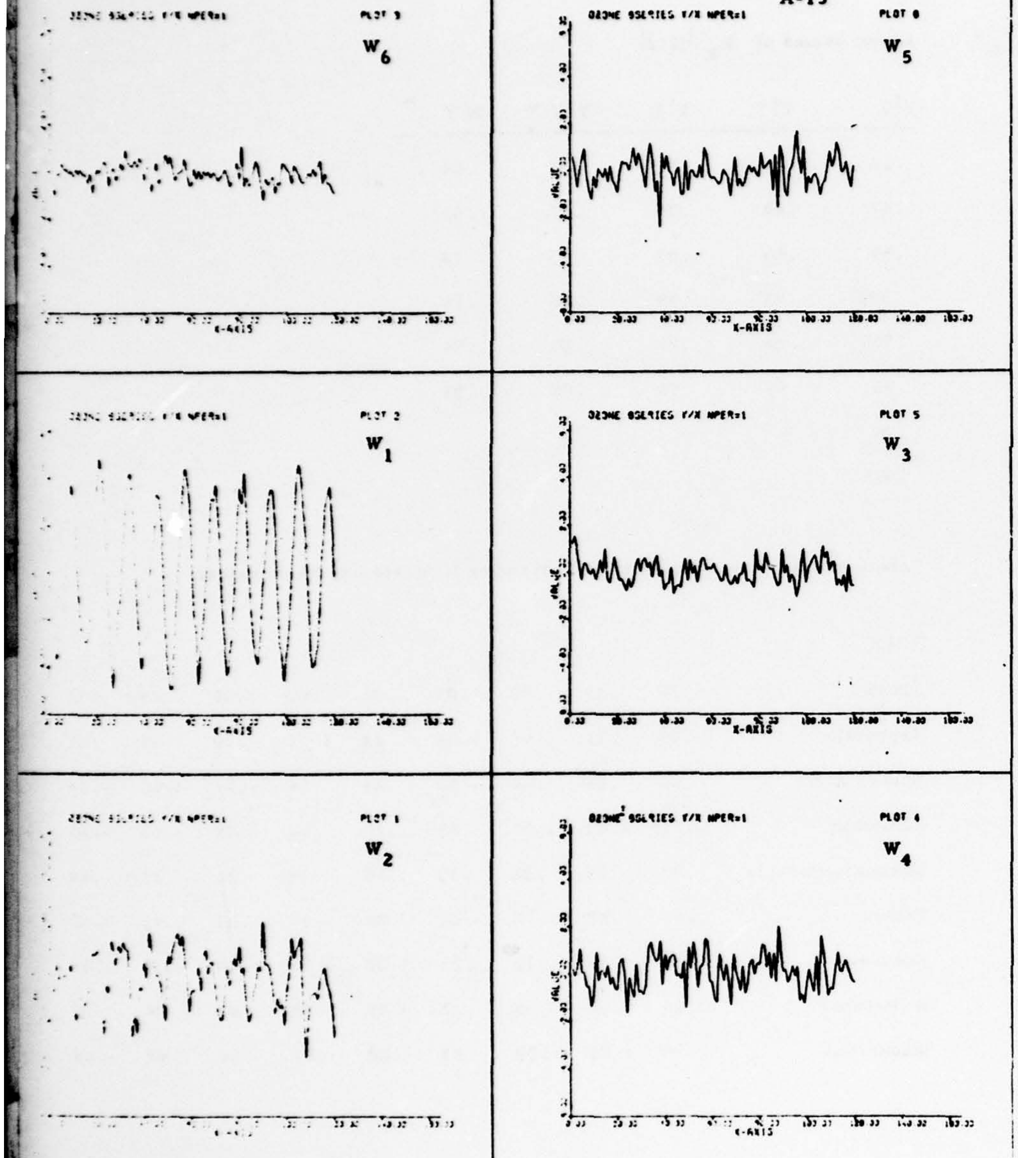



OZONE 9SERIES Y/Y.X NPER=1
PLOT 3
W4
PLOT 6
W5
PLOT 2
W1
PLOT 5
W3
PLOT 1
W2
PLOT 4
W6
VALUE
X-AXIS

A-11


OZONE 9SERIES Y/Y.X NPER=1
PLOT 3
W6
PLOT 2
W1
PLOT 1
W2
VALUE
X-AXIS

A-12



MONTHLY MEAN AND VARIANCE ADJUSTMENT

$R_Z(0)$

Arosa	1.14							
Aspendale	-.02	.75						
Buenos Aires	-.08	.16	1.00					
Edmonton	.18	-.00	.02	.92				
Macquarie Islands	-.07	.06	.08	-.00	1.06			
Tateno	.04	.00	.09	.10	.087	.88		
Huancayo	-.00	-.06	-.05	-.23	-.21	-.15	.91	
Kodaikanal	.15	-.20	-.09	-.03	-.12	-.03	.44	.81
Mauna Loa	.12	-.05	-.16	-.16	-.25	.15	.40	.24

Order of AR model fitted to Z is 1 .

$\Sigma_{Z|Z}$

Arosa	.89							
Aspendale	.04	.53						
Buenos Aires	-.11	.10	.86					
Edmonton	.10	-.01	-.07	.75				
Macquarie Islands	-.08	-.06	.04	-.07	.84			
Tateno	-.12	.00	-.01	.03	.06	.65		
Huancayo	.01	.02	-.01	-.10	-.10	-.14	.57	
Kodaikanal	-.04	.00	.00	.03	-.06	-.03	.14	.15
Mauna Loa	-.03	.02	-.11	.00	-.11	.00	.20	.04

$\Sigma_{Y'Y}$

Arosa	1.00					
Aspendale	.00	.56				
Buenos Aires	-.13	.11	.87			
Edmonton	.13	-.03	-.08	.81		
Macquarie Islands	-.07	-.07	.04	-.03	.86	
Tateno	-.11	.01	-.01	.02	.05	.66

$\Sigma_{Y'X}$

Arosa	.95					
Aspendale	.06	.66				
Buenos Aires	-.04	.10	.84			
Edmonton	.17	-.02	-.03	.76		
Macquarie Islands	-.07	.06	.04	-.07	.94	
Tateno	.00	.00	.10	.08	.11	.76

$\Sigma_{Y|Y,X}$ = 6^{th} principal minor of $\Sigma_{Z|Z}$

$\Sigma_{Y|Y,X^+}$

Arosa	.86					
Aspendale	.04	.53				
Buenos Aires	-.11	.10	.84			
Edmonton	.12	-.01	-.07	.72		
Macquarie Islands	-.10	-.06	.02	-.06	.79	
Tateno	-.11	.01	-.01	.00	.04	.61

Latent values of $R_Y^{-1}(0)\Sigma$

Z|Z	Y|Y	Y|X	Y|Y,X	Y|Y,X*
.15	.61	.71	.57	.54
.47	.64	.79	.62	.61
.59	.86	.84	.83	.78
.702	.93	.89	.86	.84
.705	.98	.95	.94	.91
.93	.99	.96	.98	.93
.96				
.985				

Latent vectors (column vector) corresponding to increasing latent values

$\Sigma_{Z|Z}$

Arosa	.12	.36	.90	-.09	-.06	.02	1.00	.25	-.31
Aspendale	-.08	-.11	.59	1.00	.24	1.00	.20	-.40	.4[illegible]
Buenos Aires	.00	.26	.60	-.33	.21	.64	-.41	1.00	-.14
Edmonton	-.08	-.09	1.00	-.33	.10	.44	-.48	-.85	-.81
Macquarie Islands	.05	-.06	.93	.59	.15	-.95	-.26	.23	-.64
Tateno	.03	.52	.78	-.32	.38	-.63	.02	-.43	1.0[illegible]
Huancayo	-.18	-.10	-.78	-.32	1.00	-.38	.45	-.33	-.56
Kodaikanal	1.00	-.48	.58	.32	-.18	.48	-.40	.04	.52
Mauna Loa	.08	1.00	-.59	.43	-.28	.23	-.54	.02	-.57

$\Sigma_{Y|Y}$

Arosa	.59	-.15	-.52	.29	.52	1.00
Aspendale	.13	1.00	-.09	1.00	.14	-.32
Buenos Aires	.53	-.09	.86	.01	-.85	.74
Edmonton	.33	.13	1.00	-.33	1.00	-.43
Macquarie Islands	.24	.63	-.42	-.95	-.05	.29
Tateno	1.00	-.30	-.52	.10	-.37	-.98

$\Sigma_{Y'X}$

Arosa	-.75	1.00	.08	.44	.18	.59
Aspendale	.54	-.99	.36	.02	.69	1.00
Buenos Aires	.71	.28	-.09	1.00	-.16	-.33
Edmonton	1.00	.69	.53	-.45	-.66	.18
Macquarie Islands	.41	.90	-.22	-.29	1.00	-.25
Tateno	-.58	-.34	1.00	.08	.36	-.56

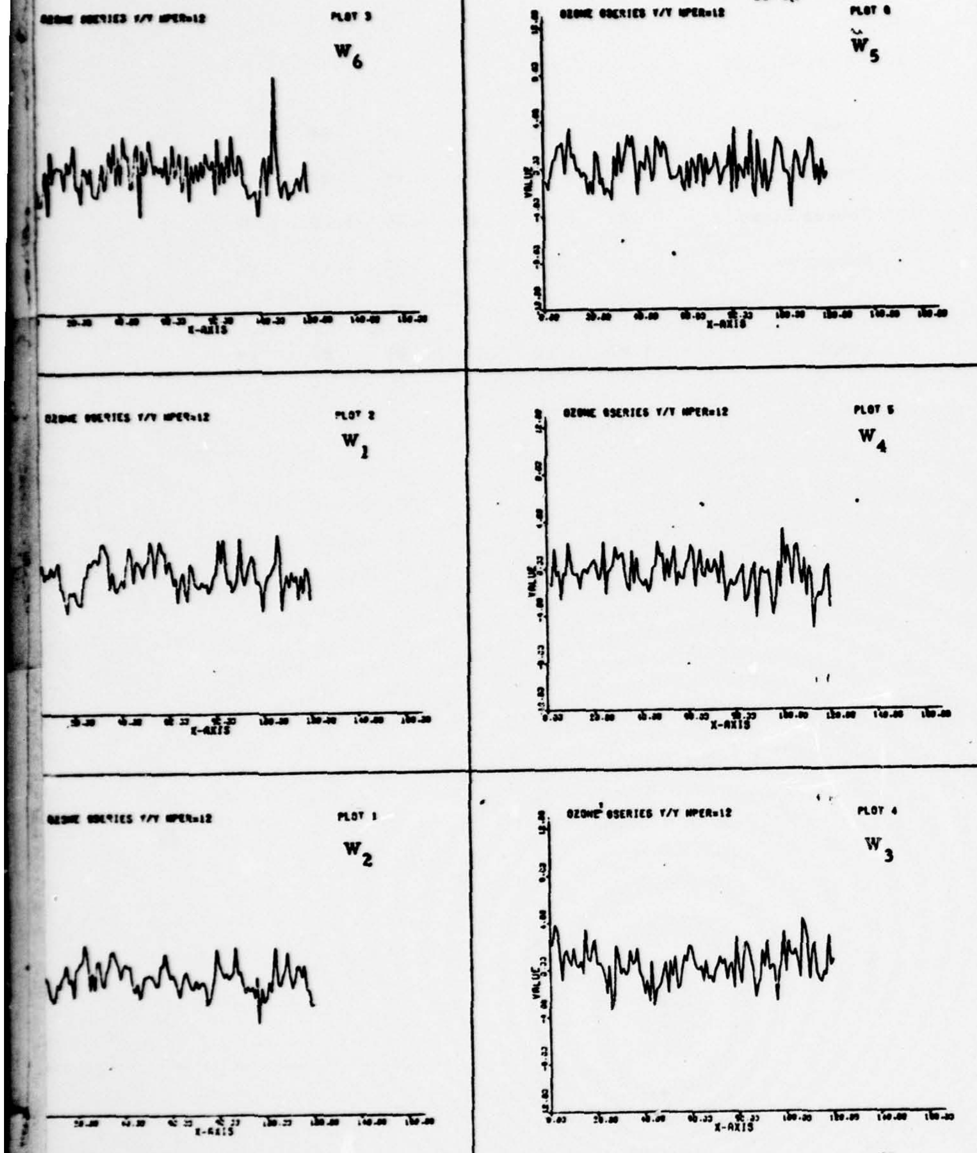
$\Sigma_{Y|Y,X}$

Arosa	.81	-.11	-.44	.56	1.00	.36
Aspendale	-.31	1.00	-.48	-.63	.97	-.34
Buenos Aires	.48	.09	.50	-.65	-.03	1.00
Edmonton	.31	.32	1.00	.31	.41	-.63
Macquarie Islands	.08	.64	-.10	1.00	-.87	.30
Tateno	1.00	-.02	-.41	-.66	-.93	-.58

$\Sigma_{Y|Y,X}$

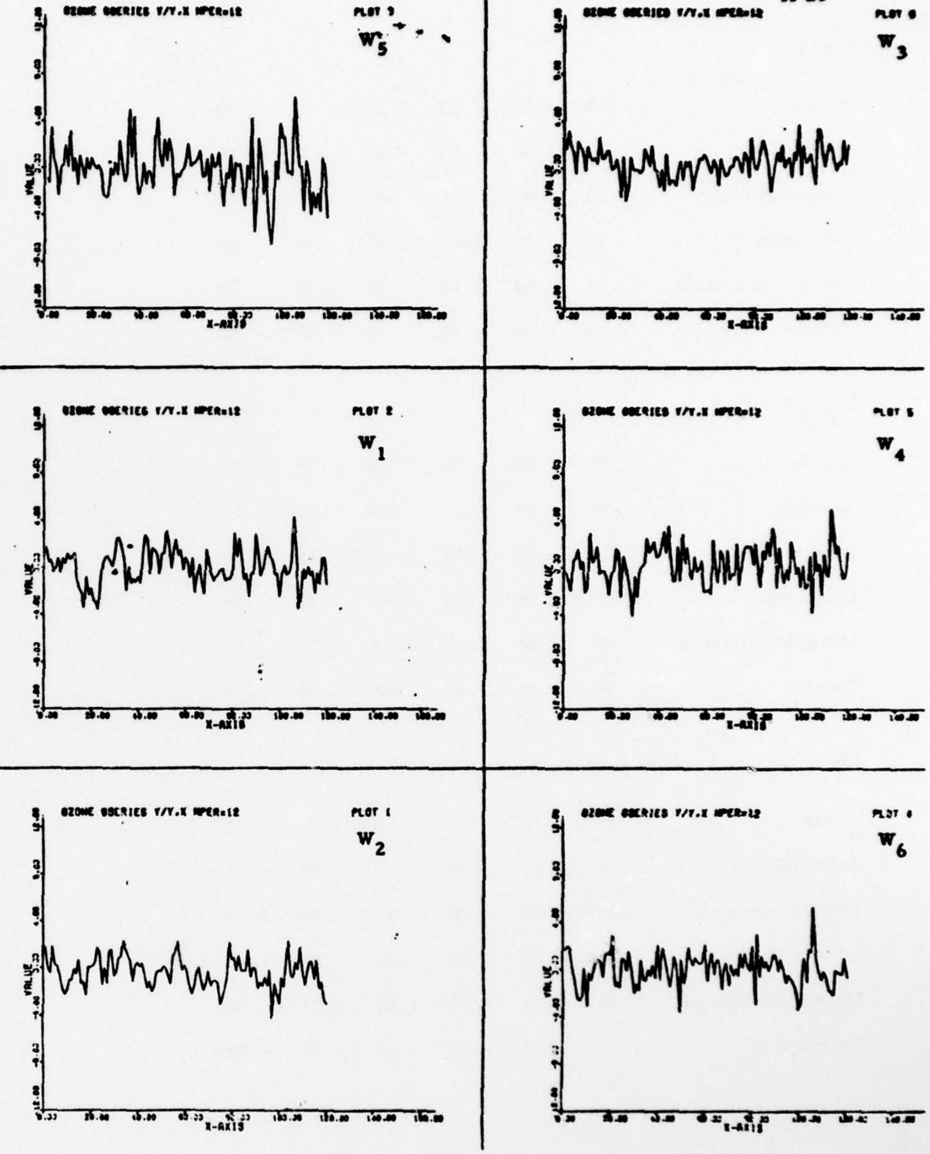
Arosa	.71	-.26	-.66	.34	.68	-.51
Aspendale	-.10	1.00	-.28	-.99	.22	-.85
Buenos Aires	.47	.01	.43	-.20	1.00	1.00
Edmonton	.41	.26	1.00	.57	-.14	-.82
Macquarie Islands	.32	.65	-.36	1.00	-.33	.58
Tateno	1.00	-.20	-.11	-.93	-.88	.19

A-19



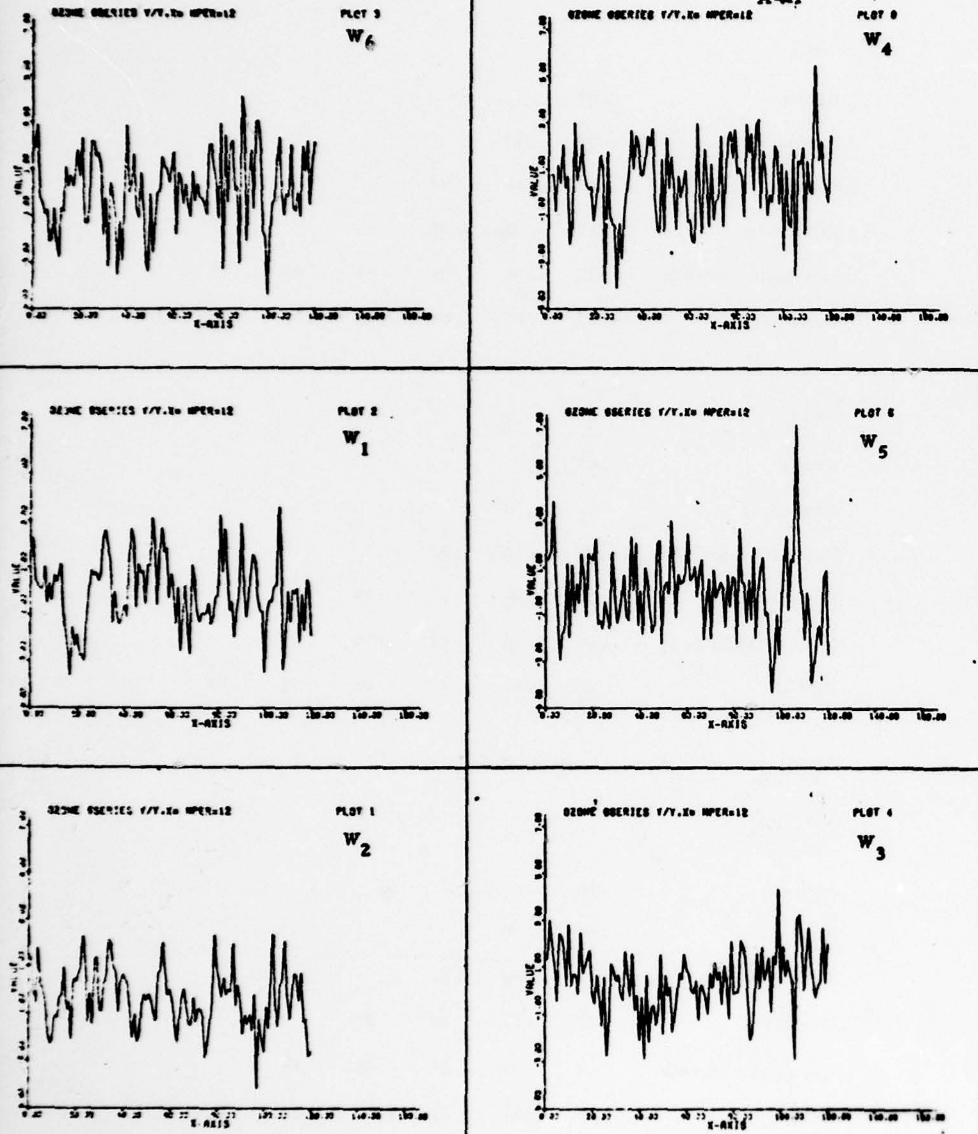
OZONE OSERIES Y/Y NPER=12
PLOT 3
W_6
PLOT 0
$\tilde{W}_5$
PLOT 2
W_1
PLOT 5
W_4
PLOT 1
W_2
PLOT 4
W_3
VALUE
X-AXIS

A-20

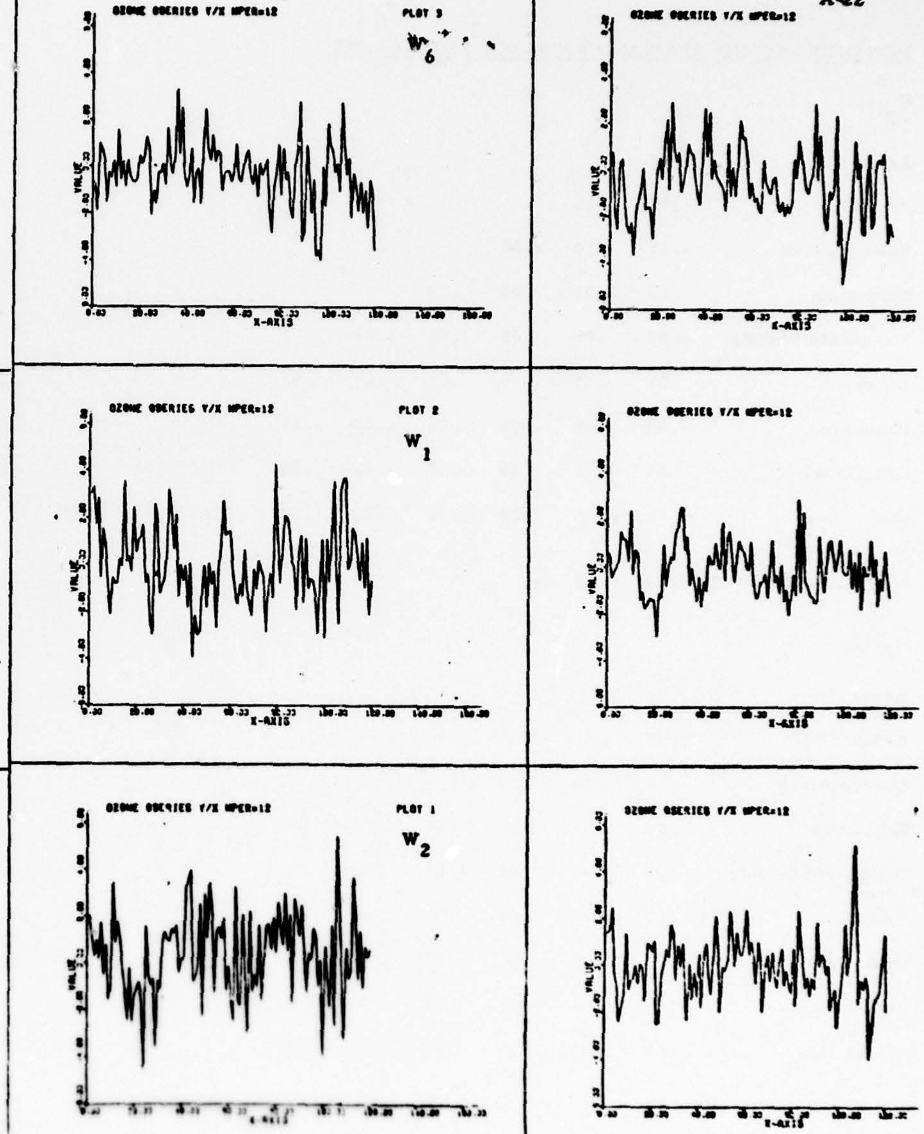


OZONE OSERIES Y/Y.X NPER=12
PLOT 3
W_5
PLOT 0
W_3
PLOT 2
W_1
PLOT 5
W_4
PLOT 1
W_2
PLOT 4
W_6
VALUE
X-AXIS

A-21


PLOT 3
W6
X-AXIS
PLOT 0
W4
PLOT 2
W1
PLOT 6
W5
PLOT 1
W2
PLOT 4
W3
VALUE

A-22


OZONE OSERIES Y/X NPER=12
PLOT 3
W6
PLOT 2
W1
PLOT 1
W2
X-AXIS
VALUE

MONTHLY FITTED MEAN AND VARIANCE ADJUSTMENT

$R_Z(0)$

Arosa	1.18								
Aspendale	-.02	.76							
Buenos Aires	-.11	.14	1.00						
Edmonton	.18	-.01	.09	.93					
Macquarie Islands	-.07	.06	.06	.02	1.07				
Tateno	.01	.02	.10	.02	.08	.91			
Huancayo	.00	-.02	-.05	-.22	-.20	-.12	.92		
Kodaikanal	.14	-.21	-.10	-.05	-.12	-.04	.41	.81	
Mauna Loa	.16	-.03	-.20	-.16	-.23	.07	.39	.21	1.10

Order of AR model fitted to Z is 1 .

$\Sigma_{Z'Z}$

Arosa	.96								
Aspendale	.02	.60							
Buenos Aires	-.16	.10	.87						
Edmonton	.10	-.01	-.01	.74					
Macquarie Islands	-.05	-.04	.01	-.04	.88				
Tateno	-.10	-.01	.00	-.02	.04	.74			
Huancayo	.02	.07	-.01	-.08	-.10	-.09	.61		
Kodaikanal	-.01	-.04	-.02	.02	-.06	-.03	.11	.17	
Mauna Lao	.05	.02	-.17	-.01	-.09	-.07	.22	.06	.66

$\Sigma_{Y|Y}$

Arosa	1.07					
Aspendale	.00	.63				
Buenos Aires	-.17	.11	.87			
Edmonton	.11	-.02	-.02	.81		
Macquarie Islands	-.05	-.04	.01	.00	.90	
Tateno	-.07	.01	.01	-.05	.02	.80

$\Sigma_{Y|X}$

Arosa	1.01					
Aspendale	.03	.67				
Buenos Aires	-.06	.10	.85			
Edmonton	.16	-.01	.07	.79		
Macquarie Islands	-.06	.06	.03	-.05	.95	
Tateno	-.05	.01	.10	.02	.11	.77

$\Sigma_{Y|Y,X}$ = 6^{th} principal minor of $\Sigma_{Z|Z}$

$\Sigma_{Y|Y,X^+}$

Arosa	.96					
Aspendale	.02	.58				
Buenos Aires	-.14	.10	.82			
Edmonton	.11	.01	-.01	.72		
Macquarie Islands	-.04	-.04	-.01	-.04	.85	
Tateno	-.10	-.00	-.01	0.02	.02	.72

Latent values of $R_Y^{-1}(0)\Sigma$

Z'Z	Y'Y	Y'X	Y|Y, X	Y|Y, X*
.20	.68	.76	.63	.60
.51	.75	.80	.70	.69
.62	.92	.85	.81	.79
.76	.96	.90	.91	.89
.81	.98	.94	.94	.91
.94	.99	.96	.98	.95
.97				
.98				

Latent vectors (column vector) corresponding to increasing latent values

$\Sigma_{Z|Z}$

Arosa	.07	.28	-.74	-.23	-.07	.77	.93	-.42	-.05
Aspendale	.00	-.09	-.49	1.00	-.07	1.00	-.61	.43	-.08
Buenos Aires	.00	.35	-.61	-.23	.44	.04	-.33	-.22	1.00
Edmonton	-.06	-.13	-.99	-.43	.28	-.17	-.02	1.00	-.37
Macquarie Islands	.04	-.08	-.68	.64	.14	-.70	1.00	.05	.18
Tateno	.03	.51	-.61	.13	.25	-.30	-.58	0.49	-.90
Huancayo	-.07	-.20	1.00	-.05	1.00	.25	.73	-.02	-.57
Kodaikanal	1.00	-.22	-.71	.29	-.30	-.05	-.73	.17	.25
Mauna Lao	-.01	1.00	.39	.14	-.21	-.33	.16	.63	.40

$\Sigma_{Y|Y}$

Arosa	.65	-.31	-.47	1.00	.30	-.68
Aspendale	.37	1.00	.63	.99	-.39	.40
Buenos Aires	.99	-.30	.49	-.42	-.90	-.86
Edmonton	.86	-.19	1.00	-.24	1.00	.64
Macquarie Islands	.63	.67	-.64	-.58	.67	-.50
Tateno	1.00	-.20	-.94	-.01	-.51	1.00

$\Sigma_{Y|X}$

Arosa	.60	-.83	.11	.56	.60	-.56
Aspendale	-.12	1.00	.56	-.62	1.00	-.62
Buenos Aires	-.41	.68	.56	1.00	-.23	-.19
Edmonton	-.59	-.91	.81	-.52	-.53	-.62
Macquarie Islands	-.62	-.78	.23	.13	.60	1.00
Tateno	1.00	.16	1.00	-.21	-.25	.70

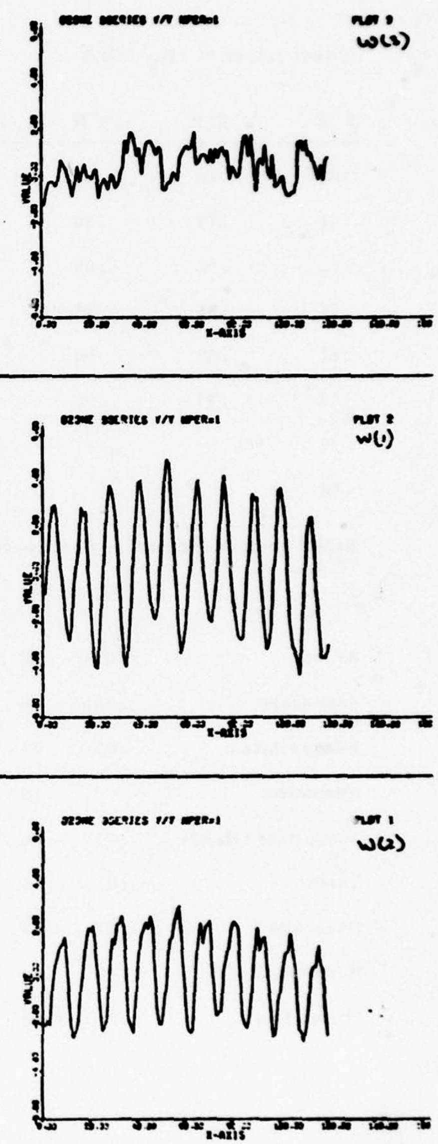
$\Sigma_{Y|Y,X}$

Arosa	.77	-.53	-.29	.91	.50	.40
Aspendale	.26	1.00	-.50	-.24	1.00	-.16
Buenos Aires	.81	-.16	.13	-.89	-.15	1.00
Edmonton	.81	.11	1.00	-.16	.28	-.67
Macquarie Islands	.49	.66	.13	1.00	-.65	.27
Tateno	1.00	-.21	-.71	-.37	-.50	-.74

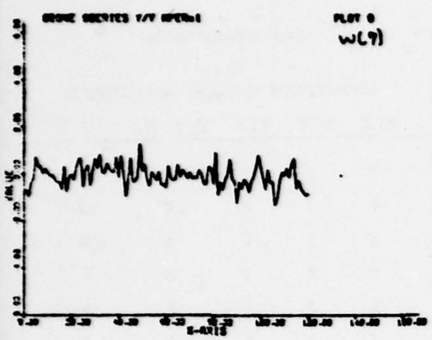
$\Sigma_{Y|Y,X^+}$

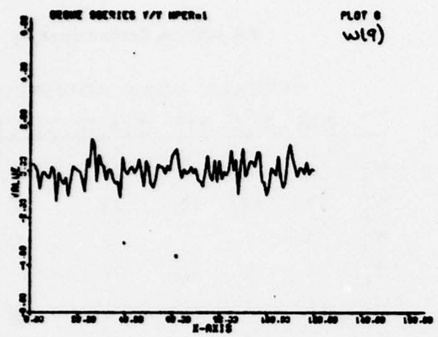
Arosa	.63	-.44	-.48	-.80	.30	-.64
Aspendale	.25	1.00	-.42	-.19	1.00	.33
Buenos Aires	.87	-.17	.04	1.00	.20	-.94
Edmonton	.77	-.12	1.00	-.23	.44	.69
Macquarie Islands	.65	.57	.26	-.53	-.84	-.49
Tateno	1.00	-.19	-.63	.24	-.48	1.00

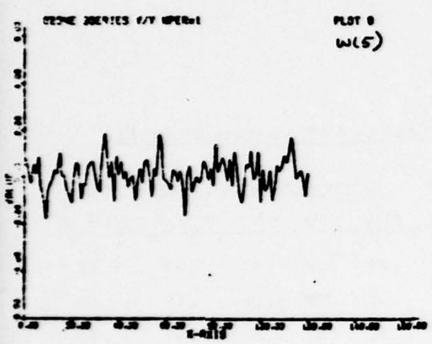
Predictable components of all nine time series, subtracting only overall mean.

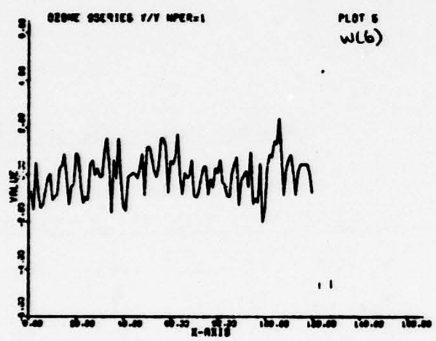


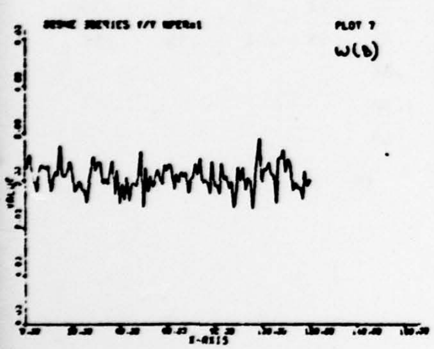
A-29

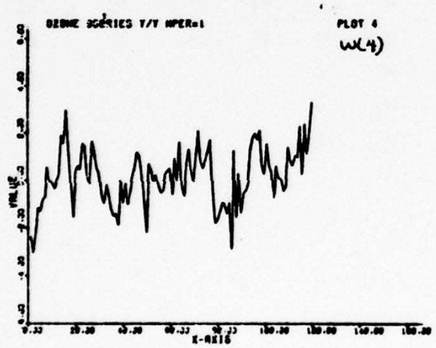






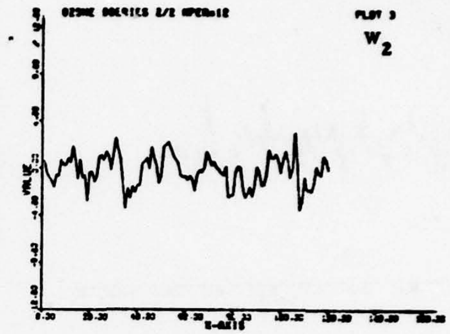


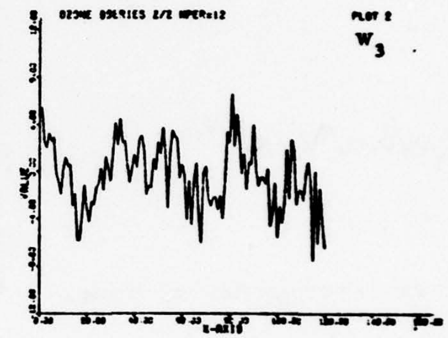


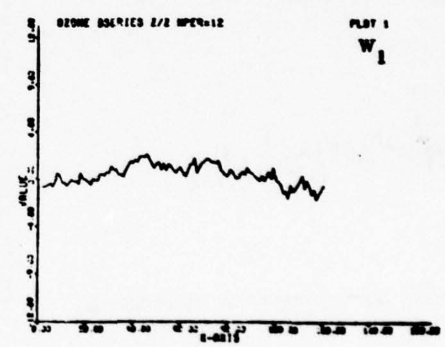


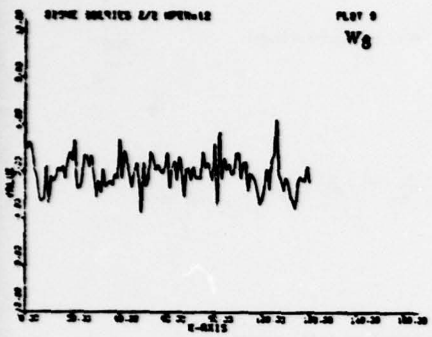
A-30

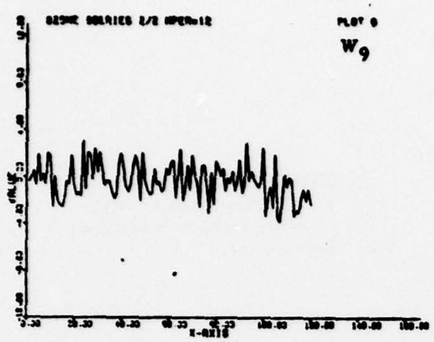
Predictable components of all nine time series; with monthly mean and variance adjustment.

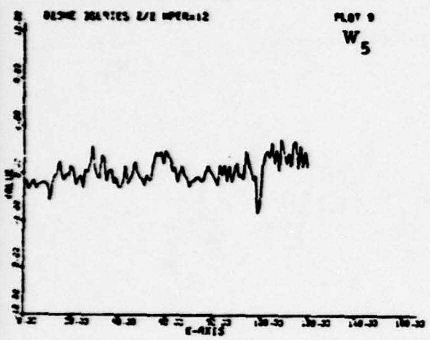


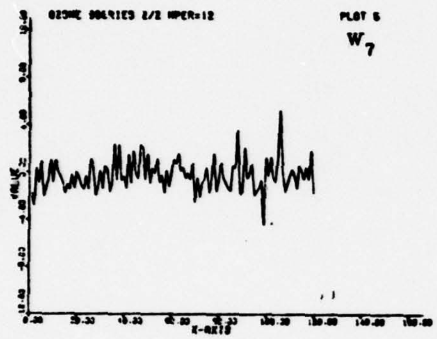


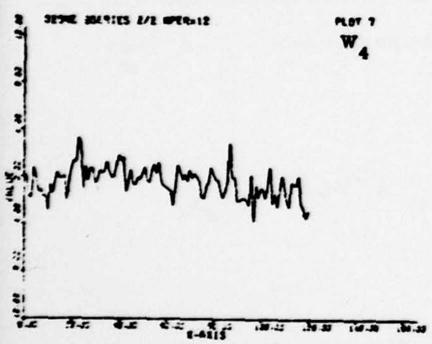


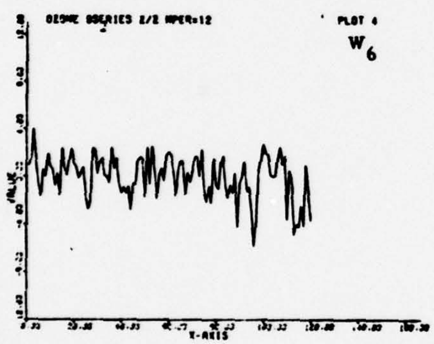












Univariate Analysis of Predictable Components

AR orders determined by CAT (W_j is j^{th} most predictable)

	OVERALL MEAN ADJUSTED					MONTHLY MEAN ADJUSTED				
	$Z \mid Z$	$Y \mid Y$	$Y \mid X$	$Y \mid Y,X$	$Y \mid Y,X^+$	$Z \mid Z$	$Y \mid Y$	$Y \mid X$	$Y \mid Y,X$	$Y \mid Y,X^+$
W_1	9	13	4	8	6	9	1	1	1	1
W_2	12	13	12	12	13	5	1	0	10	10
W_3	4	1	1	1	1	2	1	1	4	3
W_4	2	1	1	0	1	0	2	1	2	1
W_5	1	1	1	1	1	1	1	1	2	1
W_6	1	2	1	1	0	2	1	6	1	2
W_7	1					1				
W_8	7					0				
W_9	3					2				

Residual Variance for Order Determined by CAT (in units of R(0))

	OVERALL MEAN ADJUSTED					MONTHLY MEAN ADJUSTED				
	$Z \mid Z$	$Y \mid Y$	$Y \mid X$	$Y \mid Y,X$	$Y \mid Y,X^+$	$Z \mid Z$	$Y \mid Y$	$Y \mid X$	$Y \mid Y,X$	$Y \mid Y,X^+$
W_1	.07	.10	.06	.06	.06	.22	.74	.90	.77	.77
W_2	.12	.46	.47	.46	.46	.58	.77	1.00	.68	.68
W_3	.53	.87	.98	.92	.88	.71	.97	.81	.91	.93
W_4	.67	.87	.96	1.00	.94	1.00	.95	.88	.94	.95
W_5	.95	.93	.93	.90	.93	.76	.91	.94	.91	.93
W_6	.94	.95	.89	.94	1.00	.91	.95	.72	.95	.92
W_7	.94					.85				
W_8	.78					1.00				
W_9	.93					.69				